

*Citation for published version:*

Scheller, M, Garcia, S, Bathelt, J, De Haan, M & Petrini, K 2019, 'Active touch facilitates object size perception in children but not adults: A multisensory event related potential study', *Brain Research*, vol. 1723, 146381, pp. 1-9. <https://doi.org/10.1016/j.brainres.2019.146381>

*DOI:*

[10.1016/j.brainres.2019.146381](https://doi.org/10.1016/j.brainres.2019.146381)

*Publication date:*

2019

*Document Version*

Peer reviewed version

[Link to publication](#)

*Publisher Rights*

CC BY-NC-ND

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Active touch facilitates object size perception in children but not adults: a multisensory event  
related potential study

Meike Scheller<sup>a^</sup>, Sara Garcia<sup>b^</sup>, Joe Bathelt<sup>cd</sup>, Michelle de Haan<sup>d\*</sup>, Karin Petrini<sup>a\*</sup>

<sup>a</sup> Department of Psychology, University of Bath, UK; <sup>b</sup> UCL Institute of Ophthalmology; <sup>c</sup> Brain &  
Cognition, University of Amsterdam, Netherlands; <sup>d</sup> UCL Great Ormond Street Institute of Child  
Health, UK

<sup>^</sup>MS and SG are joint first authors, \*KP and MDH are joint senior authors on this work.

**Corresponding author:**

Meike Scheller,  
Department of Psychology  
University of Bath  
Claverton Down, BA2 7AY  
Bath, United Kingdom  
e-mail: m.scheller@bath.ac.uk

Declarations of interest: none

**Abstract**

In order to increase perceptual precision the adult brain dynamically combines redundant information from different senses depending on their reliability. During object size estimation, for example, visual, auditory and haptic information can be integrated to increase the precision of the final size estimate. Young children, however, do not integrate sensory information optimally and instead rely on active touch. Whether this early haptic dominance is reflected in age-related differences in neural mechanisms and whether it is driven by changes in bottom-up perceptual or top-down attentional processes has not yet been investigated. Here, we recorded event-related-potentials from a group of adults and children aged 5-7 years during an object size perception task using auditory, visual and haptic information. Multisensory information was presented either congruently (conveying the same information) or incongruently (conflicting information). No behavioral responses were required from participants. When haptic size information was available via actively tapping the objects, response amplitudes in the mid-parietal area were significantly reduced by information congruency in children but not in adults between 190ms-250ms and 310ms-370ms. These findings indicate that during object's size perception only children's brain activity is modulated by active touch supporting a neural maturational shift from sensory dominance in early childhood to optimal multisensory benefit in adulthood.

*Keywords:* Sensory dominance, active touch, multisensory, event-related potentials, congruency, development

## **Active touch facilitates object size perception in children but not adults: a multisensory event related potential study**

### **1. Introduction**

Evidence suggests that the integration of multisensory information benefits perception by providing increased precision and accuracy in a multitude of everyday tasks (e.g. from object recognition to way finding; Ernst & Banks, 2002; Landy, Banks, & Knill, 2011). Contemporary work on the development of this process, however, has shown that children up until the age of 10-12 years, do not use multisensory information in the same way adults do, but rather rely dominantly on the sense that is more robust for the task at hand (Adams, 2016; Gori, Del Viva, Sandini, & Burr, 2008; Nardini, Jones, Bedford, & Braddick, 2008; Petrini, Remark, Smith, & Nardini, 2014). For example, Gori and colleagues (2008) asked adults and children aged 5-10 years old children to discriminate between the orientation and size of objects using either touch, vision, or both at the same time. They found that, while adults integrated haptic and visual information in a statistically optimal fashion, children focused predominantly on one sense while almost completely ignoring information from the other one (Gori et al., 2008). That is, during orientation discrimination children focused mostly on vision while size discrimination was dominated by the active touch percept. This haptic dominance during size discrimination has further been replicated using auditory and haptic cues (Petrini et al., 2014). Nevertheless, the reasons for the late development of adult-like multisensory integration and the long persistence of haptic dominance in childhood is poorly understood. Understanding how the temporal processing of ecologically relevant information changes throughout childhood can provide insights into the extent to which multisensory perceptual mechanisms rely on top-down attentional control or the developmental maturation of low-level perceptual neural properties.

Up to now, few studies have examined the neural mechanisms of multisensory processing during development (e.g. Brandwein et al., 2011; Brett-Green, Miller, Gavin, & Davies, 2008; Johanssen & Röder, 2014) and fewer still have focused on somatosensory-auditory integration (Brett-Green et al., 2008). For example, using an event related potential (ERP) paradigm, Brandwein et al (2011) investigated the developmental trajectory of neural processing of audio-visual information in children aged 7-16 years old. They found that, behaviourally, children's levels of integration gradually changed and reached mature levels at around 15 years of age, while the neural correlates for this integration could already be detected at around 10-11 years old. A systematic relationship between age and the brain processes underlying audio-visual integration was revealed in the auditory N1 ERP component (~120 ms), with audio-visual peak amplitude changing from being more positive than the sum of visual and auditory amplitudes in 7-9 year-olds to being more negative than the sum of the visual and auditory amplitude in adults. In contrast, Brett-Green et al. (2008) found somatosensory-auditory integration effects in the signal amplitude of the P1-N1-P2 ERP complex in children aged 6-13 years old, similar to what has previously been reported in adults (Foxe et al., 2000; Murray et al., 2005). In their sample they did, however, not differentiate between age-groups of children, thus leaving unclear whether there were developmental changes between 6 and 13 years old.

These developmental studies, as well as the few auditory-somatosensory studies in adults, suggest that the neurophysiological benefit driven by multisensory integration can be detected already at early stages of sensory processing (<150ms), which is characteristic of low-level or bottom-up mechanisms (De Meo, Murray, Clarke, & Matusz, 2015). This evidence, however, has been obtained in either audio-visual studies (e.g. Brandwein et al., 2011; De Meo et al., 2015) or studies using passive tactile stimuli (e.g. Brett-Green et al., 2008; Foxe et al., 2000; Sperdin, Cappe, Foxe, & Murray, 2009) so it is yet to be examined at which processing

stage haptic dominance arises. Indeed, active touch requires cognitive control and attentional resources due to its goal-directed nature (De Meo et al., 2015) and thus could be largely influenced by top-down processes and detectable only at a later stage of processing.

Given the consistent results showing a dominance of active touch over both sound and vision during object size discrimination in childhood (Gori et al., 2008; Petrini et al., 2014), this study focuses on the mechanisms of haptic dominance using an ERP paradigm to measure sensory processing. Our investigation focuses specifically on mid-parietal regions because this area is known to play an important role in the integration of multisensory information. Previous findings have shown that both children and adults exhibit differences in the P1-N1-P2 components between simultaneous auditory-somatosensory responses and summed unisensory responses in midline and central/post-central scalp regions (Brett-Green et al., 2008; Foxe et al., 2000; Murray et al., 2005). Furthermore, consistent developmental differences in topographical parietal regions in ERP multisensory studies have previously been shown (Brandwein et al., 2011; Johanssen & Röder, 2014). These consistent developmental differences over parietal channels may indicate a critical role of the IPS, which has been identified as one of the areas classically associated with multisensory processing of tactile, auditory and visual information (e.g. Ben Hamed, Duhamel, Bremmer, & Graf, 2001; Saito, Okada, Morita, Yonekura, & Sadato, 2003; Stilla & Sathian, 2008).

To detect the modulatory effect of active touch on object size perception, we adapted the task of Petrini et al. (2014) and presented two balls of different size in either audio-only, audio-visual, or audio-visual-haptic condition while manipulating the congruency between these different sensory modalities. That is, we used multisensory incongruency as an indicator of object size consistency among the senses and predicted that children would show a larger difference in ERP amplitude than adults for congruent and incongruent haptic information, given children's higher reliance on active touch. We expected to find a larger effect of haptic

congruency in children when compared to adults given the commonly found effect of multisensory integration (e.g. Brandwein et al., 2011; Brett-Green et al., 2008; Foxe et al., 2000; Murray et al., 2005). We further hypothesized that we would find a differential activation in the N2 component, given that our paradigm introduced a multisensory conflict, which has repeatedly been shown to modulate this component in adults (Forster & Pavone, 2008; Gu, Mai, & Luo, 2013; Lindström, Paavilainen, Kujala, & Tervaniemi, 2012; Lu, Ho, Sun, Johnson, & Thompson, 2016).

## 2. Results

Figure 1 shows the grand average ERPs for audio, audio-visual and audio-visual-haptic conditions for mid-parietal channels (61, 62, 67, 72, 77, 78). Children showed overall higher responses than adults. Difference waves between responses to congruent and incongruent stimuli were calculated for both, adults and children (see left panels in Figure 2). Please see the supplemental material for a component-based analysis returning similar results.

*Display figure 1 about here.*

ERP difference waves based on comparison between congruent and incongruent trials showed a significant deviation from 0 for latencies between 190ms-250ms ( $p < .05$  for 16 consecutive time points) and 310ms-370ms ( $p < .05$  for 16 consecutive time points). However, this was only true for children in the audio-visual-haptic condition and not for the audio-visual condition, or for adults (see figure 2). Scalp topography plots show activity to audio-visual-haptic stimulus presentation occurred primarily in posterior areas.

*Display Figure 2 about here.*

### 3. Discussion

Combining different senses can reduce uncertainty in everyday tasks and thus improve our precision. Whilst adults can integrate different sensory information optimally by making a weighted linear average based on the reliability of each sensory modality, children younger than 8 years cannot. For example, children do not integrate active touch with either sound or vision optimally when perceiving and judging objects' sizes, but show a strong dominance of the haptic information (Gori et al., 2008; Petrini et al., 2014). Whereas previous studies have examined the development of multisensory integration in childhood using electrophysiological methods and passive tactile stimulation (e.g. Brandwein et al., 2011; Brett-Green et al., 2008), it is not yet known how sensory dominance of active touch is reflected in neurophysiological differences between young children and adults. Furthermore, whilst evidence suggests that supra- or super-additive multisensory integration can occur already at early stages of processing (<150ms; e.g. Brett-Green et al., 2008; De Meo et al., 2015; Foxe et al., 2000; Sperdin et al., 2009), it is unknown whether sensory dominance arises at such early stages of processing as well.

Our results revealed a reduction in amplitude during time windows that are temporally aligned with the N2 (190ms-250ms) and the P3b (310ms-370ms) component in mid-parietal regions for children but not for adults. However, this was only the case when haptic congruent information about the object size was available. Interestingly, this reduction in amplitude was not evident when congruent auditory and visual information were presented alone.

We predicted a modulation of the N2 component based on several EEG/ERP studies reporting its sensitivity towards multisensory conflict. For example, response amplitudes to multisensory stimuli that were incongruent in spatial location, temporal synchrony, direction of movement, or emotional content have been found to be larger with respect to congruent pairings (Forster & Pavone, 2008; Gu et al., 2013; Lindström et al., 2012; Lu et al., 2016;



Nieuwenhuis et al., 2003). The findings of the current study confirm this directionality of amplitude modulation as children showed a larger amplitude in the audio-visual-haptic incongruent than congruent condition in a time window and channel region that are consistent with the N2 component. Neural processing changes in the parietal region in children may indicate a critical role of the IPS, which has been consistently associated with multisensory processing (Bolognini & Maravita, 2007; Kamke, Vieth, Cottrell, & Mattingley, 2012; Sereno & Huang, 2014). Furthermore, IPS has been shown to have a crucial role in global object perception (Amedi, von Kriegstein, van Atteveldt, Beauchamp, & Naumer, 2005; Bodegård, Geyer, Grefkes, Zilles, & Roland, 2001; Faillenot, Decety, & Jeannerod, 1999; Roland, O'Sullivan, & Kawashima, 1998; Sathian et al., 2011) and the on-line adjustment of monitoring and adjusting grasping movements to object size (Glover, Miall, & Rushworth, 2005). However, we cannot be sure about the exact generators, which should be investigated in a future study with a larger sample size.

In contrast to children, adults did not show a congruency-dependent modulation of N2. This result may indicate that children younger than 8 years have a higher sensitivity to multisensory conflict when active touch is involved as compared to adults. The reason why adults do not show such modulation may result from a dominance of auditory information when judging object size that is reflected by a significantly higher precision in sound discrimination (Petrini et al., 2014). As shown in previous studies assessing attentional modulation of cross-modal interactions in adults, touch, in comparison to vision or audition, possesses a special role because it can be decoupled from the other senses (Eimer, Van Velzen, & Driver, 2002). That is, while auditory and visual cueing influenced sensory processing in the respective other sense, touch did not bias sensory processing in either visual (Eimer & Driver, 2000) nor auditory (Eimer et al., 2002) attended cues in adult participants. Hence, as the ability to use sound is greater in adults than children when estimating object size, adults may have ignored or

197 decoupled haptic information while children may have been unable to ignore such information.  
198 This is also supported by research showing that children are often unable to ignore irrelevant  
199 sensory information (Innes-Brown et al., 2011; Petrini, Jones, Smith, & Nardini, 2015).

200 We further found a significant difference between responses to congruent and  
201 incongruent stimuli in children in the audio-visual-haptic condition during a later time window  
202 (310ms-370ms) that we did not predict. The latencies and parietal distribution response  
203 difference are associated with the conflict-sensitive P3b, a subcomponent of the P300  
204 component (Polich, 2004). This subcomponent has been shown to be involved in memory  
205 processing (Polich, 2007) and visuo-motor learning (Morgan, Luu, & Tucker, 2016). It has  
206 further been shown to be modulated by levels of expertise (Morgan et al., 2016) which could  
207 explain why, in the present study, we found differences between children and adults in this  
208 component.

209 Contrary to our expectation, we did not find any age-related modulation of the N1 when  
210 active touch was involved. This suggests that haptic dominance in children does not arise at  
211 earlier stages (<150ms) of processing but rather at later stages (De Meo et al., 2015) and might  
212 therefore be modulated by top-down processes. This difference between our results and those  
213 of studies reporting earlier occurrences of somatosensory-auditory integration (e.g. Brett-  
214 Green et al., 2008; Foxe et al., 2000; Sperdin et al., 2009) may be explained by the use of active  
215 touch instead of passive tactile stimulation. Given that active touch is usually goal-directed and  
216 thus requires motor planning and attentional control, latency differences between  
217 somatosensory and haptic information processing would not be surprising. Therefore, this later  
218 influence of haptic dominance on sensory processing may not be generalizable to other  
219 multisensory dominance processes that do not involve goal-directed actions. ERP  
220 developmental studies of action monitoring and cognitive control do show a modulation of N2  
221 component for congruent and incongruent stimuli (e.g. Ladouceur, Dahl, Carter, 2004, 2007;

Buss et al., 2011) similarly to our study, with incongruent stimuli inducing a larger N2 response than congruent stimuli. However, these studies show an increase in N2 or N2 difference waves response with age rather than a decrease in this component as we have shown here. Hence, our results cannot be fully explained by age-related changes in cognitive control or action monitoring. A possible explanation of why our results differ from those of the aforementioned studies is that our participants did not performed a task as we did not want to assess the effect of error monitoring performance in children and adults (e.g. Ladouceur, Dahl, Carter, 2004, 2007; Buss et al., 2011) but rather the age-related effect of active touch on object perception. Future research could further examine whether the involvement of top-down control over sensory dominance is essential or whether sensory dominance as assessed with passive multisensory stimulation is a purely perceptual mechanism.

Our results are in agreement with the few multisensory developmental studies using neurophysiological measures that show a change in multisensory processing from early childhood to adulthood. In the study by Brandwein and colleagues (2011), a clear developmental change in response amplitude was shown within a time window ~100ms-140ms as well as in the parietal region between 190ms and 240ms (see Figure 9 in Brandwein et al. 2011). The amplitude of the audio-visual response was less negative than the sum of the auditory and visual responses in younger children (7-9-year-old), had the same level of negativity for the 10-12-year-old children, and became more negative than the auditory and visual sum for the older groups (13-16-year-old children and adults). This prolonged maturation of adult-like multisensory processing during childhood appears to be led by a prioritization of the unisensory process (Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006; Stein, Labos, & Kruger, 1973; Stein & Meredith, 1993; Wallace & Stein, 2001), which is not surprising given the many physical and physiological changes occurring in this developmental period (e.g. changes in structural and functional brain organization, physical

changes in hand and body size, and maturation of the auditory cortex; Moore & Linthicum, 2007; Paus, 2005; Ponton, Eggermont, Kwong, & Don, 2000; Steinberg, 2005). In other words, before adult-like multisensory integration can be fully achieved, the individual senses need to stabilize through a cross-calibration process requiring the most reliable sense to teach the less reliable sense to accurately process the perceptual properties of objects in the environment (Gori et al., 2008). In the current study, we support this view by showing that in children, in contrast to adults, haptic information (known to dominate children's object size perception by calibrating vision and auditory information for object size judgment; Gori et al., 2008; Petrini et al., 2014) modulates the brain response elicited by the visual and auditory information. Our support to the cross-calibration hypothesis (Gori et al., 2008), however, seem to sit in contrast to recent findings showing a lack of motorsensory recalibration in children aged 8-11 years (Vercillo et al., 2014). While the development of recalibration (shift after adaptation to sensory asynchrony) and cross-calibration are probably linked, they are not the same mechanism as cross-calibration as discussed in our paper is a process by which the most robust sense for one task teaches the other senses so that accurate perception can be achieved. In Vercillo's et al., (2014) study children did not show motorsensory recalibration probably because children have been shown by several studies to have lower sensitivity to multisensory asynchrony (i.e. bigger temporal integration/binding window: e.g. Stanley et al., 2019; Hillock-Dunn & Wallace, 2012; Hillock, Powers, & Wallace, 2011), that is, children, compared to adults, keep perceiving synchrony for larger delays between the senses and consequently may not recalibrate because they do not perceive the delay. In our study there was no motorsensory asynchrony, the only manipulation was in sensory congruency, which has been shown to affect children's as well as adults' performance (Petrini et al., 2014) of the same age tested in our study. Hence, while children may gain from cross-calibration, i.e. gain from the inclusion of the most robust sense

for the task when perceiving an object, they may not recalibrate due to their lower sensitivity to asynchrony.

An alternative explanation of the age-related effect found here could refer to the mere differences in number of sensory cues in the audio-visual and audio-visual-haptic conditions, i.e. one condition has two sensory cues while the other has three. However, it is difficult to explain these findings based merely on the number of senses involved given that in adults the trimodal and bimodal conditions give rise to the same ERP results. This lack of difference between trimodal and bimodal conditions in the current study for the adult group stands in contrast to evidence showing that in adults trimodal conditions generate a multisensory advantage when compared to bimodal and unimodal conditions (e.g. Diederich & Colonius, 2004). Moreover, the effect is specific to the N2 component in children. If the number of senses was the main driver of the findings then it is unclear why both the audio-visual and audio-visual-haptic condition led to a decrease in amplitude for earlier components in children, despite the different number of cues (see supplemental material). Thus, the dominance of haptic information for size perception in childhood appears to be a more plausible explanation for the findings. Nevertheless, future neurophysiological studies with adult participants could compare different bimodal combinations (audio-visual, audio-haptic, and visual-haptic) to a more naturalistic, multimodal (here trimodal) setting (e.g. audio-visual-haptic).

Despite our best efforts to match the visual information in AV and AVH these two visual conditions were not exactly the same. The only way they could have been matched perfectly while still keeping the task realistic (i.e. rather than using recorded videos for AV condition) was through immersive virtual reality. Unfortunately using the EEG cap with the head mounted display would have been problematic and would have added more noise to the data. However, given the progress made to integrate these technologies in recent years (e.g. Ehinger et al., 2014) it should be possible to overcome this limitation in future studies using a similar

paradigm. Nevertheless, the component-based analysis in the supplemental material does show similar results for AV and AVH in P2 which point to a similarity in processing of the two conditions. Finally, although the number of retained trials per participant was low, a recent paper has shown that the internal consistency of event-related potentials associated with cognitive control in N2 and P3 can be achieved after 14-20 trials (Rietdijk et al., 2014). This together with the fact that both the component-based and difference waves analysis returned the same results indicate that our results are reliable despite the low number of trials.

### *Conclusion*

Based on behavioural evidence it has been established that young children do not integrate active touch with either sound or vision when perceiving and judging objects' size but rather show a strong dominance of the haptic information (Gori et al., 2008; Petrini et al., 2014). To the best of our knowledge, however, the underlying changes in brain activity that may reflect this haptic dominance in early childhood have not been previously documented. Here, we used electrophysiology to examine the difference in brain processes between young children and adults when using active touch to perceive and judge objects' size. Our ERP data show a modulation of the neural response during two distinct time-windows that are aligned with the conflict-sensitive N2 and P3b components. These, however, are only showing in response to haptic in/congruent information in children, but not adults. This result is consistent with the behavioral data showing overreliance of children on haptic information aimed at facilitating the cross-sensory calibration needed to achieve an adult-like multisensory mechanism of object perception. This study is the first to use active touch in an ERP study to assess sensory dominance in young children and its results are an important benchmark against which to assess the development of this sensory developmental stage in different clinical and sensory impaired

populations with known atypical multisensory processing (e.g. autism, schizophrenia, and individuals with visual, hearing, or motor impairments).

## 4. Methods

### 4.1 Participants

10 children aged between 5 and 7 years ( $M = 6$ ,  $SD = 0.82$ ; 4 female) and 10 adults aged between 20 and 31 years ( $M = 23.9$ ,  $SD = 3.57$ ; 6 female) participated. The age range for the recruited children was based on the behavioral evidence that children younger than 8 years do not integrate haptic information with either vision and sound but instead show a strong dominance of the haptic information when perceiving and judging object size (Gori et al., 2008; Petrini et al., 2014). The number of participants is similar to the previous developmental ERP study using passive tactile stimulation to investigate somatosensory-auditory integration (Brett-Green et al., 2008). There was one left handed person in the adult group and two in the group of children as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The remaining participants were either ambidextrous with a preference for using the right hand or fully right-handed. Ethics permission was granted by the UCL Great Ormond Street Institute of Child Health Research ethics committee (02CN01).

### 4.2 Stimuli

Visual and haptic stimuli were two wooden balls (see Figure 1S in the supplemental material), a big ball (57mm diameter) and a small ball (41mm diameter). The corresponding sound amplitude for the big ball was 79dB and for the small ball was 71dB. The stimuli for the visual, haptic and auditory modalities were selected from a set of balls and corresponding sounds previously used and tested by Petrini et al. (2014). The two ball and corresponding

sound sizes used here were at the ends of the range used in Petrini et al. (2014) and were easily perceived as different in size by two separate samples of children and adults of the same age of the participants taking part to this study (see Figure 2 in Petrini et al., 2014). The reason for selecting two balls and sounds that could be easily differentiated based on size (i.e. through differences in height for touch and loudness for sound) is so that we could be sure all participants perceived the difference in size with touch and sound without the need of asking for behavioral responses. We opt not to ask for behavioral responses based on a relevant previous developmental study examining the integration of somatosensory and sound information (Brett-Green et al., 2008). Sound recordings lasted for 66ms.

#### 4.3 Procedure

Participants were seated comfortably and asked to rest their dominant hand on a rectangular (5cm deep) semi-soft foam surface, covering a (Touch ProE-X, Keytech Inc Magic, Texas, USA) touch screen, positioned before them on a table. A speaker (Logitech, Lausanne, Switzerland) was positioned directly below the touch screen, underneath the table, hidden from view (see Figure 3). Randomization of condition order and stimuli presentation were controlled using Matlab (Version R2014a, The MathWorks Inc., Natick, MA) and the Psychophysics toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997), on a Dell computer running Windows XP Service Pack 3 (Microsoft Inc., NW).

*Display figure 3 about here.*

Before the experiment started participants were shown the two balls and they all easily identified the smaller and bigger ball. Participants completed five blocks (see Table 1) of 48 trials with a short break after each set of 6 trials. The trials were started manually as in Petrini



et al. (2014). That means that the time between trials was variable. We used unpredictable interval to minimise expectation effects as suggested by Woodman (2010) to avoid that the alpha-wave activity of participants became phase locked with the stimulus presentation rate. Also, we triggered the onset of the next trial manually to ensure that all participants were paying attention when the stimulus was presented. During the experiment, participants were asked to look at a square hole that had been cut in the middle of the rectangular foam surface on which their hand was rested. The experimenter monitored the participants attention by making sure that participants kept looking at the square hole at all time during the block. The experimenter (but not the participant) was able to look at changes in looking behavior thanks to a video recording of the participant shown on the computer. During ‘audio’ blocks, participants were asked to listen to the ball sounds that were triggered by the experimenter pressing a mouse button. Here, the only information about object size that participants received was auditory (either the sound of the small ball or of the big ball). During ‘audio-visual’ blocks, the experimenter placed one of the wooden balls within the square hole, and participants were asked to observe the experimenter tap the ball. The tap elicited pressure on the touch screen which triggered the sound to be played back instantaneously. During ‘audio-visual-haptic’ blocks, the experimenter placed the ball inside the square hole and asked the participant to tap the ball themselves. Thereby, participants were instructed to keep their hand straight and flat during the tapping movement in order to minimize any effects that might result from differences in hand sizes between adults and children (as in Petrini et al., 2014). That is, as adults have larger hands, curving the hand would allow them to assess more information than children. Keeping the hand straight and flat was supposed to control for such differences. Once the ball was tapped and pressure was sensed on the touch screen, the sound recording was played. This sound was either congruent (e.g. 41mm ball and 71dB sound) or incongruent (e.g. 41mm ball and 79dB sound) with the size information delivered by haptic and visual

modalities. The visual information (the visible action of tapping the ball) was kept as similar as possible in the audio-visual and audio-visual-haptic condition by positioning the experimenter on the same side of the arm used by the participant. A thin 1cm layer of soft foam was inserted between the thick layer of foam and the touch screen positioned underneath to eliminate any impact sound between the wooden balls and the hard surface of the touch screen). The visual stimulation persisted longer than the sound of the ball hitting the surface. Participants were instructed to close their eyes while the experimenter was positioning the ball in the square hole of the soft foam surface. Participants were then allowed to look at the ball and either tap it themselves (audio-visual-haptic condition) or observe the experimenter tapping the ball (audio-visual condition). The sound was triggered by the haptic tap without time delays.

*Display table 1 about here.*

#### *4.4 Data acquisition*

The EEG was recorded using a GES 200 high-density, high-impedance recording system, a NetAmps 200 amplifier and HydroCel Geodesic Sensor Nets with 128 channels (Electrical Geodesics Inc., OR). Recordings were obtained using NetStation software V4.1.2 (Electrical Geodesics Inc., OR) on an Apple Macintosh PowerPC G4 running Mac OS 10.3.9. The sample frequency was set to 250Hz. A vertex reference was used for recording.

The amplifier was calibrated and impedances were measured before each recording. Channels with impedance higher than 50k $\Omega$  were checked for good contact with the scalp and adjusted where necessary. An event code was sent to NetStation via Matlab whenever an auditory stimulus was triggered.

#### 414 4.5 EEG Processing and Analysis

415 Due to an error during recording (a failed communication between NetStation and  
416 Matlab), the EEG was not recorded for trials in which the big ball was tapped (the data for  
417 these trials were not saved), therefore only the EEG for trials in which the small ball was tapped  
418 with either the small ball (congruent condition) or big ball (incongruent condition) sound were  
419 included in the analyses. Not including the data for the tapped big ball did not affect the data  
420 for the tapped small ball as these two conditions were carried out separately. The EEG signal  
421 was filtered off-line using a 0.1Hz high-pass filter and 30Hz low-pass filter. The ERP is  
422 dominated by lower EEG frequencies. Higher frequencies are less relevant for the ERP and  
423 may be contaminated by high frequency noise, such as muscle artefacts. A low-pass filter at  
424 30Hz is therefore routinely applied across many studies. The high-pass filter was applied to  
425 remove low-frequency drift associated with electrode shearing or drying, and 0.1 Hz is what is  
426 recommended when testing children (Acunzo et al., 2012). Channels with weak correlation ( $r$   
427  $\leq 0.3$ ) to neighboring channels were removed. Timing tests indicated a 24ms delay between  
428 stimulus trigger (haptic tap) and sound presentation. ERPs were time-locked to the sound. This  
429 delay was systematically adjusted for every logged event presentation. EEG recordings were  
430 then segmented into 650ms epochs, starting at -50ms before the trigger until 600ms following  
431 the trigger. For each segment, channels with a peak-to-peak amplitude larger than 80 $\mu$ v were  
432 replaced using spherical interpolation of neighboring channel values using EEGLAB v13.2.1  
433 functions. The spherical interpolation is the recommended and default approach to replace  
434 noisy electrodes in EEGLAB. This method provides the most accurate interpolation for high-  
435 density EEG (Perrin et al. 1989). Any segments for which 30% of all channels exceeded 100 $\mu$ v  
436 were excluded. The EEG was re-referenced to the average reference, with eye channels being  
437 excluded. Baseline correction was applied for a short temporal window of 10ms-20ms post-  
438 stimulus. This temporal window was chosen in order to avoid incorporating movement artifacts

that result from arm movements during the audio-visual-haptic condition into the baseline correction and is in line with previous research investigating multisensory integration of audio-somatosensory stimuli (Foxy et al., 2000). Segmented trials were submitted to an Independent Component Analysis (ICA). A kurtosis threshold of 3 was used to detect unusually peaky activity distributions which are likely to represent artifacts (Delorme, Sejnowski, & Makeig, 2007), and were consequently removed. Following artifact rejection, an average was computed for each participant for each condition. Children retained, on average, 12 audio-big, 18 audio-small, 14 audio-visual-incongruent, 14 audio-visual-congruent, 14 audio-visual-haptic-incongruent, and 14 audio-visual-haptic-congruent trials. Adults retained, on average, 22 audio-big, 24 audio-small, 24 audio-visual-incongruent, 24 audio-visual-congruent, 23 audio-visual-haptic-incongruent and 23 audio-visual-haptic-congruent trials.

Individual artifact-free trials were combined to individual averages for each condition, upon which grand-average waveforms for each condition were computed, including all participants in both age groups. Difference waves based on the congruency of the stimuli were calculated for each multisensory condition and each individual by subtracting responses to congruent stimuli from responses to incongruent stimuli. Only the size information of the modality of interest (e.g. haptic) differed between the congruent and incongruent condition. Grand averages of the difference waves were computed for both the adult and children group. As we used three different modes of stimulus presentation (audio, audio + visual, and audio + visual + active touch) we could not directly compare between these conditions. This is because the three conditions differ in the amount of muscle movement (noise). In order to allow for a comparison between these conditions, we made use of the sensory congruency effect by testing congruent vs incongruent stimulus pairs, and then looked at these congruency effects in the three conditions. As the amount of movement was different between the A, AV, AVH conditions, but not between congruent and incongruent stimulus pairs within each of these

conditions, we were able to compare whether the neural response to the congruency between the two stimuli differs when touch is available (AVH), compared to when is not available (AV).

The difference waves reported here are simply the subtracted potential of incongruent from congruent responses. Basically, what this allows us to do is to subtract the activity from other processes that are similar between congruent/incongruent conditions (e.g. noise due to arm movements/motor planning activity in AVH), and hence are not related to perceiving the object's size based on congruent multisensory information. For direct comparison between similar conditions (that only vary by one factor, in our case congruency), difference waves calculation is recommended (Kappenman & Luck, 2016).

In order to identify the temporal onset of processing differences, time windows that showed significantly different responses between congruent and incongruent information processing were defined. For that purpose, the difference potentials were segmented into 163 discrete time points across the whole duration of the segment from -50ms to 600ms. For each time point a paired comparison (paired t-test) of the difference wave amplitude and 0 was conducted. In order to control for an inflation of alpha error, we estimated the sequence length that was necessary to indicate significant deviations from 0. That is, deviations from 0 would only be deemed significant if a sequence of twelve or more consecutive time points (~46ms) would yield statistically significant results (see Guthrie & Buchwald, 1991). This method is an alternative to a traditional correction for multiple comparisons and has been developed by Guthrie and Buchwald (1991) for difference potentials analysis (the analysis used here). Alpha inflation is taken into account by not accepting every significant test result, unless 12 or more significant results occur in a row. As the probability of finding 12 consecutive significant t-test results is low (for 160 time points/testes), we only consider a signal significant once it exceeds 12 consecutive significant t-test results. The amount of consecutive time points (i.e. 12 in our

case) depends on the amount of comparisons made per difference waves (i.e. how many p-values are compared), autocorrelation in the data, and the number of subjects. We calculated the amount of consecutive time points for both adults and children separately. This method has been used by previous multisensory studies (e.g. Stekelenburg & Vroomen, 2007; Butler et al., 2012). Please see supplemental material for a component-based analysis of the data.

We focused on mid-parietal channels based on previous findings showing alterations in early-latency signals between simultaneous auditory-somatosensory responses and summed unisensory responses at midline and central/post-central scalp regions (Brett-Green et al., 2008; Foxe et al., 2000; Murray et al., 2005). Furthermore, several multisensory ERP studies have shown consistent developmental differences in the parietal region (Brandwein et al., 2011; Johanssen & Röder, 2014). The EEG, however, was recorded with standard nets that contained 128 electrodes. The high density of electrodes has several advantages that were exploited in the current analysis. First, all electrodes were used for the analysis to calculate the average reference. This is the closest possible approximation to a reference-free recording (Nunez & Srinivasan, 2006). Second, each position in the 10-20 system is covered by several electrodes so that an average signal can be calculated that is less influenced by noise that may affect individual electrodes. Further, the high-density of electrodes provides better spatial information about the topography of the ERP. The main statistical comparison focused on channel regions that were indicated based on previous studies. The other electrodes were not excluded but were not of primary interest for the analysis.

**509 Acknowledgements**

510           We would like to thank the children and their parents or guardians for their  
511 participation.

**512 Funding**

513           This work was supported by the British Academy/Leverhulme [grant number:  
514 SG142127]. M.S. was supported by a University of Bath Humanities and Social Sciences PhD  
515 studentship. J.B. was supported by a UCL Impact/Great Ormond Street Children's Charity PhD  
516 studentship (11CN06) and by an MRC post-doctoral fellowship (5PQ40). M.d.H. was  
517 supported by Great Ormond Street Hospital Children's Charity.

518

## References

- Acunzo, D. J., MacKenzie, G., and van Rossum, M. C. W. (2012). Systematic biases in early ERP and ERF components as a result of high-pass filtering. *J. Neurosci. Methods* 209, 212–218. doi: 10.1016/j.jneumeth.2012.06.011
- Adams, W. J. (2016). The development of audio-visual integration for temporal judgements. *PLoS Computational Biology*, 12(4), e1004865. <https://doi.org/10.1371/journal.pcbi.1004865>
- Amedi, A., von Kriegstein, K., van Atteveldt, N. M., Beauchamp, M. S., & Naumer, M. J. (2005). Functional imaging of human crossmodal identification and object recognition. *Experimental Brain Research*, 166(3–4), 559–571. <https://doi.org/10.1007/s00221-005-2396-5>
- Ben Hamed, S., Duhamel, J.-R., Bremmer, F., & Graf, W. (2001). Representation of the visual field in the lateral intraparietal area of macaque monkeys: a quantitative receptive field analysis. *Experimental Brain Research*, 140(2), 127–144. <https://doi.org/10.1007/s002210100785>
- Bodegård, A., Geyer, S., Grefkes, C., Zilles, K., & Roland, P. E. (2001). Hierarchical processing of tactile shape in the human brain. *Neuron*, 31(2), 317–328. [https://doi.org/10.1016/S0896-6273\(01\)00362-2](https://doi.org/10.1016/S0896-6273(01)00362-2)
- Bolognini, N., & Maravita, A. (2007). Proprioceptive alignment of visual and somatosensory maps in the posterior parietal cortex. *Current Biology*, 17(21), 1890–1895. <https://doi.org/10.1016/j.cub.2007.09.057>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.



<https://doi.org/10.1163/156856897X00357>

Brandwein, A. B., Foxe, J. J., Russo, N. N., Altschuler, T. S., Gomes, H., & Molholm, S. (2011). The development of audiovisual multisensory integration across childhood and early adolescence: A high-density electrical mapping study. *Cerebral Cortex*, 21(5), 1042–1055. <https://doi.org/10.1093/cercor/bhq170>

Brett-Green, B. A., Miller, L. J., Gavin, W. J., & Davies, P. L. (2008). Multisensory integration in children: A preliminary ERP study. *Brain Research*, 1242, 283–290. <https://doi.org/10.1016/j.brainres.2008.03.090>

Buss, K. A., Dennis, T. A., Brooker, R., Sippel, L. M. (2011). An ERP study of conflict monitoring in 4–8-year old children: associations with temperament. *Developmental Cognitive Neuroscience*, 1 (131–140), 131–140.

Butler, J. S., Foxe, J. J., Fiebelkorn, I. C., Mercier, M. R., Molholm, S. (2012). Multisensory representation of frequency across audition and touch: high density electrical mapping reveals early sensory-perceptual coupling. *J Neurosci*, 32, 15338–15344.

De Meo, R., Murray, M. M., Clarke, S., & Matusz, P. J. (2015). Top-down control and early multisensory processes: chicken vs. egg. *Frontiers in Integrative Neuroscience*, 9, 17. <https://doi.org/10.3389/fnint.2015.00017>

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, 34(4), 1443–1449. <https://doi.org/10.1016/j.neuroimage.2006.11.004>

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*,

- 563        66(8), 1388–1404. <https://doi.org/10.3758/BF03195006>
- 564    Ehinger, B. V., Fischer, P., Gert, A. L., Kaufhold, L., Weber, F., Pipa, G., et al. (2014).  
565        Kinesthetic and vestibular information modulate alpha activity during spatial navigation:  
566        a mobile EEG study. *Front. Hum. Neurosci.* 8:71. doi: 10.3389/fnhum.2014.00071
- 567    Eimer, M., & Driver, J. (2000). An event-related brain potential study of cross-modal links in  
568        spatial attention between vision and touch. *Psychophysiology*, 37(5), 697–705.  
569        <https://doi.org/10.1017/S0048577200990899>
- 570    Eimer, M., Van Velzen, J., & Driver, J. (2002). Cross-modal interactions between audition,  
571        touch, and vision in endogenous spatial attention: ERP evidence on preparatory states  
572        and sensory modulations. *Journal of Cognitive Neuroscience*, 14(2), 254–271.  
573        <https://doi.org/10.1162/089892902317236885>
- 574    Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a  
575        statistically optimal fashion. *Nature*, 415(6870), 429–433.  
576        <https://doi.org/10.1038/415429a>
- 577    Faillenot, I., Decety, J., & Jeannerod, M. (1999). Human brain activity related to the  
578        perception of spatial features of objects. *NeuroImage*, 10(2), 114–124.  
579        <https://doi.org/10.1006/nimg.1999.0449>
- 580    Forster, B., & Pavone, E. F. (2008). Electrophysiological correlates of crossmodal visual  
581        distractor congruency effects: Evidence for response conflict. *Cognitive, Affective and*  
582        *Behavioral Neuroscience*, 8(1), 65–73. <https://doi.org/10.3758/CABN.8.1.65>
- 583    Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E.  
584        (2000). Multisensory auditory-somatosensory interactions in early cortical processing

- 585 revealed by high-density electrical mapping. *Cognitive Brain Research*, 10(1–2), 77–83.  
586 [https://doi.org/10.1016/S0926-6410\(00\)00024-0](https://doi.org/10.1016/S0926-6410(00)00024-0)
- 587 Glover, S., Miall, R. C., & Rushworth, M. F. S. (2005). Parietal rTMS disrupts the initiation  
588 but not the execution of on-line adjustments to a perturbation of object size. *Journal of*  
589 *Cognitive Neuroscience*, 17(1), 124–136. <https://doi.org/10.1162/0898929052880066>
- 590 Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate  
591 visual and haptic form information. *Curr Biol*, 18(9), 694–698.  
592 <https://doi.org/10.1016/j.cub.2008.04.036>
- 593 Gu, Y., Mai, X., & Luo, Y. jia. (2013). Do bodily expressions compete with facial  
594 expressions? Time course of integration of emotional signals from the face and the body.  
595 *PLoS ONE*, 8(7), e66762. <https://doi.org/10.1371/journal.pone.0066762>
- 596 Guthrie, D., & Buchwald, J. S. (1991). Significance Testing of Difference Potentials.  
597 *Psychophysiology*, 28(2), 240–244.
- 598 Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory  
599 temporal binding window persist into adolescence. *Dev Sci*, 15(5), 688–696.
- 600 Hillock, A. R., Powers, A. R., & Wallace, M. T. (2011). Binding of sights and sounds: age-  
601 related changes in multisensory temporal processing. *Neuropsychologia*, 49(3), 461–467.
- 602 Innes-Brown, H., Barutcu, A., Shivdasani, M. N., Crewther, D. P., Grayden, D. B., Paolini,  
603 A. G., & Paolini, A. (2011). Susceptibility to the flash-beep illusion is increased in  
604 children compared to adults. *Developmental Science*, 14(5), 1089–1099.  
605 <https://doi.org/10.1111/j.1467-7687.2011.01059.x>
- 606 Johannsen, J., & Röder, B. (2014). Uni- and crossmodal refractory period effects of event-

- 607 related potentials provide insights into the development of multisensory processing.  
608 *Frontiers in Human Neuroscience*, 8, 552. <https://doi.org/10.3389/fnhum.2014.00552>
- 609 Kamke, M. R., Vieth, H. E., Cottrell, D., & Mattingley, J. B. (2012). Parietal disruption alters  
610 audiovisual binding in the sound-induced flash illusion. *NeuroImage*, 62(3), 1334–1341.  
611 <https://doi.org/10.1016/j.neuroimage.2012.05.063>
- 612 Kappenman, E. S., & Luck, S. J. (2016). Best practices for event-related potential research in  
613 clinical populations. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*,  
614 1(2), 110–115.
- 615 Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & Cornelissen, F.  
616 (2007). What's new in Psychtoolbox-3. In *Perception* (Vol. 36, pp. 1–235).  
617 <https://doi.org/10.1177/03010066070360S101>
- 618 Ladouceur, C. D., Dahl, R. E., Carter, C. S. (2004). ERP correlates of action monitoring in  
619 adolescence. *Ann NY Acad Sci*, 1021, 329–336.
- 620 Ladouceur, C. D., Dahl, R. E., Carter, C. S. (2007). Development of action monitoring  
621 through adolescence into adulthood: ERP and source localization. *Dev Sci*, 10, 874–891.
- 622 Landy, M. S., Banks, M. S., & Knill, D. C. (2012). Ideal-Observer Models of Cue Integration.  
623 In J. Trommershäuser, K. Kording, & M. S. Landy (Eds.), *Sensory Cue Integration* (pp.  
624 5–29). Oxford University Press.  
625 <https://doi.org/10.1093/acprof:oso/9780195387247.003.0001>
- 626 Lindström, R., Paavilainen, P., Kujala, T., & Tervaniemi, M. (2012). Processing of  
627 audiovisual associations in the human brain: dependency on expectations and rule  
628 complexity. *Frontiers in Psychology*, 3, 159. <https://doi.org/10.3389/fpsyg.2012.00159>

- 629 Lu, X., Ho, H. T., Sun, Y., Johnson, B. W., & Thompson, W. F. (2016). The influence of  
630 visual information on auditory processing in individuals with congenital amusia: An  
631 ERP study. *NeuroImage*, 135, 142–151.  
632 <https://doi.org/10.1016/j.neuroimage.2016.04.043>
- 633 Moore, J. K., & Linthicum, F. H. (2007). The human auditory system: A timeline of  
634 development. *International Journal of Audiology*, 46(9), 460–478.  
635 <https://doi.org/10.1080/14992020701383019>
- 636 Morgan, K. K., Luu, P., & Tucker, D. M. (2016). Changes in P3b Latency and Amplitude  
637 Reflect Expertise Acquisition in a Football Visuomotor Learning Task. *PLOS ONE*,  
638 11(4), e0154021. <https://doi.org/10.1371/journal.pone.0154021>
- 639 Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., ...  
640 Foxe, J. J. (2005). Grabbing your ear: Rapid auditory-somatosensory multisensory  
641 interactions in low-level sensory cortices are not constrained by stimulus alignment.  
642 *Cerebral Cortex*, 15(7), 963–974. <https://doi.org/10.1093/cercor/bhh197>
- 643 Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in  
644 human navigation. *Current Biology*, 18(9), 689–693.  
645 <https://doi.org/10.1016/j.cub.2008.04.021>
- 646 Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006).  
647 Development of multisensory spatial integration and perception in humans.  
648 *Developmental Science*, 9(5), 454–464. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-7687.2006.00512.x)  
649 [7687.2006.00512.x](https://doi.org/10.1111/j.1467-7687.2006.00512.x)
- 650 Nieuwenhuis, S., Yeung, N., Van Den Wildenberg, W., & Ridderinkhof, K. R. (2003).  
651 Electrophysiological correlates of anterior cingulate function in a go/no-go task: Effects

- 652 of response conflict and trial type frequency. *Cognitive, Affective and Behavioral*  
653 *Neuroscience*, 3(1), 17–26. <https://doi.org/10.3758/CABN.3.1.17>
- 654 Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory.  
655 *Neuropsychologia*, 9(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- 656 Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence.  
657 *Trends in Cognitive Sciences*, 9(2), 60–68. <https://doi.org/10.1016/j.tics.2004.12.008>
- 658 Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming  
659 numbers into movies. *Spatial Vision*, 10(4), 437–442.  
660 <https://doi.org/10.1163/156856897X00366>
- 661 Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp  
662 potential and current density mapping. *Electroencephalography & Clinical*  
663 *Neurophysiology*, 72, 184–187.
- 664 Petrini, K., Jones, P. R., Smith, L., & Nardini, M. (2015). Hearing where the eyes see:  
665 Children use an irrelevant visual cue when localizing sounds. *Child Development*, 86(5),  
666 1449–1457. <https://doi.org/10.1111/cdev.12397>
- 667 Petrini, K., Remark, A., Smith, L., & Nardini, M. (2014). When vision is not an option:  
668 Children’s integration of auditory and haptic information is suboptimal. *Developmental*  
669 *Science*, 17(3), 376–387. <https://doi.org/10.1111/desc.12127>
- 670 Polich, J. (2004). Clinical application of the P300 event-related brain potential. *Physical*  
671 *Medicine and Rehabilitation Clinics of North America*, 15(1), 133–161. Retrieved from  
672 <http://www.ncbi.nlm.nih.gov/pubmed/15029903>
- 673 Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical*

- 674       *Neurophysiology : Official Journal of the International Federation of Clinical*  
675       *Neurophysiology*, 118(10), 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- 676   Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central  
677       auditory system activity: Evidence from multi-channel evoked potentials. *Clinical*  
678       *Neurophysiology*, 111(2), 220–236. [https://doi.org/10.1016/S1388-2457\(99\)00236-9](https://doi.org/10.1016/S1388-2457(99)00236-9)
- 679   Rietdijk, W. J., Franken, I. H., & Thurik, A. R. (2014). Internal consistency of event-related  
680       potentials associated with cognitive control: N2/P3 and ERN/Pe. *PLOS One*, 9 (7),  
681       e102672.
- 682   Roland, P. E., O’Sullivan, B., & Kawashima, R. (1998). Shape and roughness activate  
683       different somatosensory areas in the human brain. *Proceedings of the National Academy*  
684       *of Sciences*, 95(6), 3295–3300. <https://doi.org/10.1073/pnas.95.6.3295>
- 685   Saito, D. N., Okada, T., Morita, Y., Yonekura, Y., & Sadato, N. (2003). Tactile–visual cross-  
686       modal shape matching: a functional MRI study. *Cognitive Brain Research*, 17(1), 14–25.  
687       [https://doi.org/10.1016/S0926-6410\(03\)00076-4](https://doi.org/10.1016/S0926-6410(03)00076-4)
- 688   Sathian, K., Lacey, S., Stilla, R., Gibson, G. O., Deshpande, G., Hu, X., ... Glielmi, C.  
689       (2011). Dual pathways for haptic and visual perception of spatial and texture  
690       information. *NeuroImage*, 57(2), 462–475.  
691       <https://doi.org/10.1016/j.neuroimage.2011.05.001>
- 692   Serenó, M. I., & Huang, R.-S. (2014). Multisensory maps in parietal cortex. *Current Opinion*  
693       *in Neurobiology*, 24(1), 39–46. <https://doi.org/10.1016/j.conb.2013.08.014>
- 694   Sperdin, H. F., Cappe, C., Foxe, J. J., & Murray, M. M. (2009). Early, low-level auditory-  
695       somatosensory multisensory interactions impact reaction time speed. *Frontiers in*

- 696        *Integrative Neuroscience*, 3, 2. <https://doi.org/10.3389/neuro.07.002.2009>
- 697     Stanley, B. M., Chen, Y.-C., Lewis, T. L., Maurer, D., & Shore, D. I. (2019). Developmental  
698        changes in the perception of audiotactile simultaneity. *Journal of Experimental Child*  
699        *Psychology*, 183, 208–221.
- 700     Stein, B. E., Labos, E., & Kruger, L. (1973). Sequence of changes in properties of neurons of  
701        superior colliculus of the kitten during maturation. *Journal of Neurophysiology*, 36(4),  
702        667–679. <https://doi.org/10.1152/jn.1973.36.4.667>
- 703     Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT  
704        Press. <https://doi.org/10.3389/neuro.01.019.2008>
- 705     Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in*  
706        *Cognitive Sciences*, 9(2), 69–74. <https://doi.org/10.1016/j.tics.2004.12.005>
- 707     Stekelenburg J. J., Vroomen J. (2007). Neural correlates of multisensory integration of  
708        ecologically valid audiovisual events. *J. Cogn. Neurosci.*, 19 (12), 1964-1973.
- 709     Stilla, R., & Sathian, K. (2008). Selective visuo-haptic processing of shape and texture.  
710        *Human Brain Mapping*, 29(10), 1123–1138. <https://doi.org/10.1002/hbm.20456>
- 711     Vercillo, T., Burr, D., Sandini, G., Gori, M. (2014). Children do not recalibrate motor-sensory  
712        temporal order after exposure to delayed sensory feedback. *Dev Sci*, 18(5), 703–712.
- 713     Wallace, M. T., & Stein, B. E. (2001). Sensory and multisensory responses in the newborn  
714        monkey superior colliculus. *The Journal of Neuroscience*, 21(22), 8886–8894.  
715        <https://doi.org/10.1523/JNEUROSCI.21-22-08886.2001>
- 716     Woodman, G. F. (2010). A brief introduction to the use of event-related potentials (ERPs) in



717 studies of perception and attention. *Atten Percept Psychophys*, 72:2031–2046.

718

719

720

**Figure legends**

*Figure 1.* Grand average event related potentials (ERPs) for auditory (black), audio-visual (blue) and audio-visual-haptic (red) conditions, at midline-parietal channels for adults (top three panels) and children (bottom three panels). Time 0 represents the onset of the auditory stimulus. Shaded error bands around the means represent the standard error. Note that the response to the auditory stimulus is plotted in both, the congruent and incongruent response plots to serve as a reference.

*Figure 2.* Difference waves for the audio-visual and audio-visual-haptic conditions for adults (top panel) and children (lower panel). Time windows for which significant deviations from 0 exist in the audio-visual haptic condition are shaded in grey. Scalp topographies are displayed for the mean activity at mid-latency of the two time windows that were identified (190ms-250ms, 310ms-370ms) for both age groups.

*Figure 3.* Sketch of the experimental setup and the three conditions. Upper panel shows the experimenter (left) and participant (right) interacting with the stimuli in the audio-visual-haptic condition. Here, the participant was tapping the wooden ball while looking at it. A sound was played in response to the tap from a speaker positioned out of view from right underneath the setup. Lower panel shows a sketch of the three conditions including the participant listening to the sound alone (audio), listening to the sound while observing the experimenter tapping the ball (audio-visual), and the sound paired with seeing and tapping the ball themselves (audio-visual-haptic).

Figures

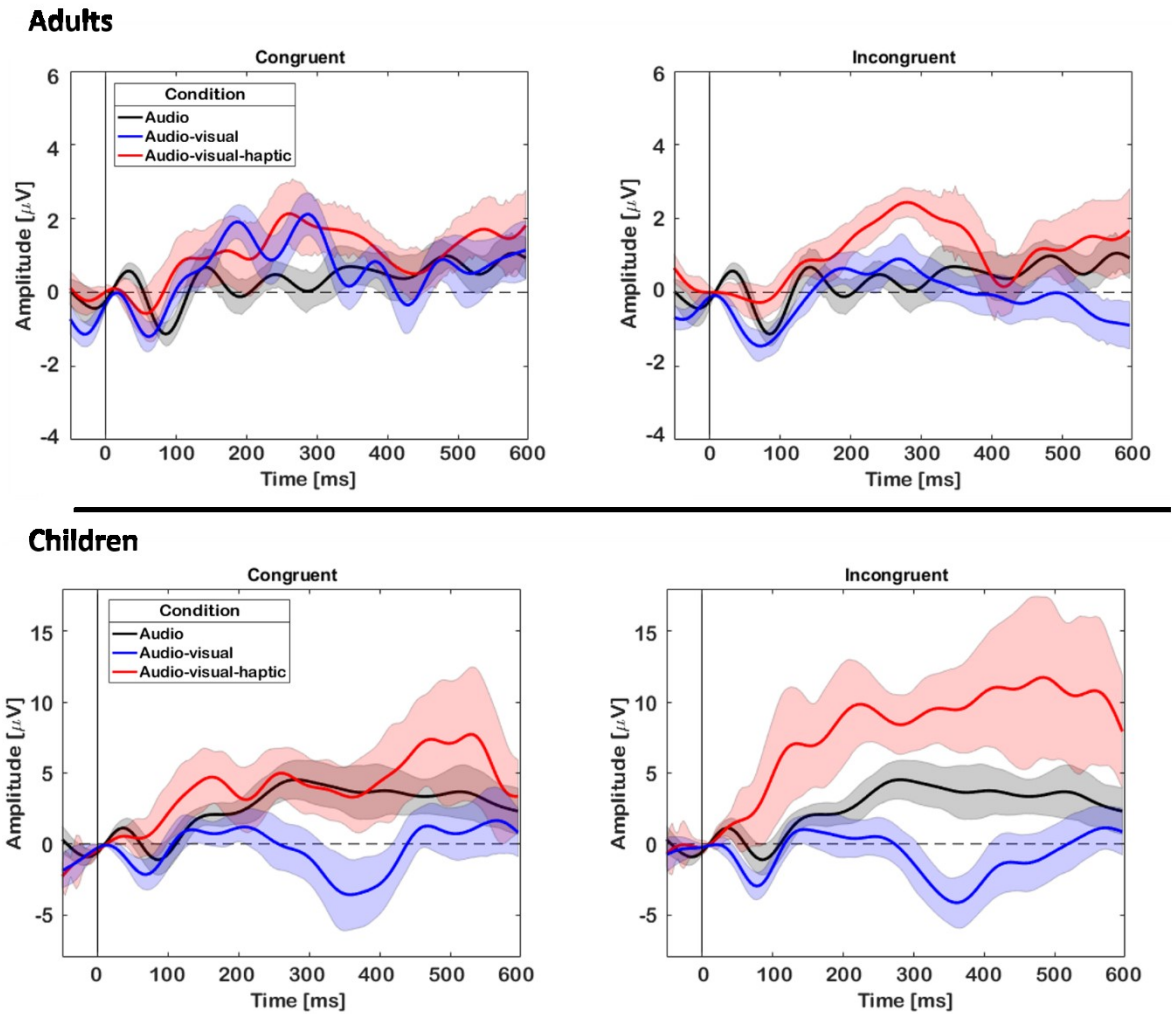


Figure 1

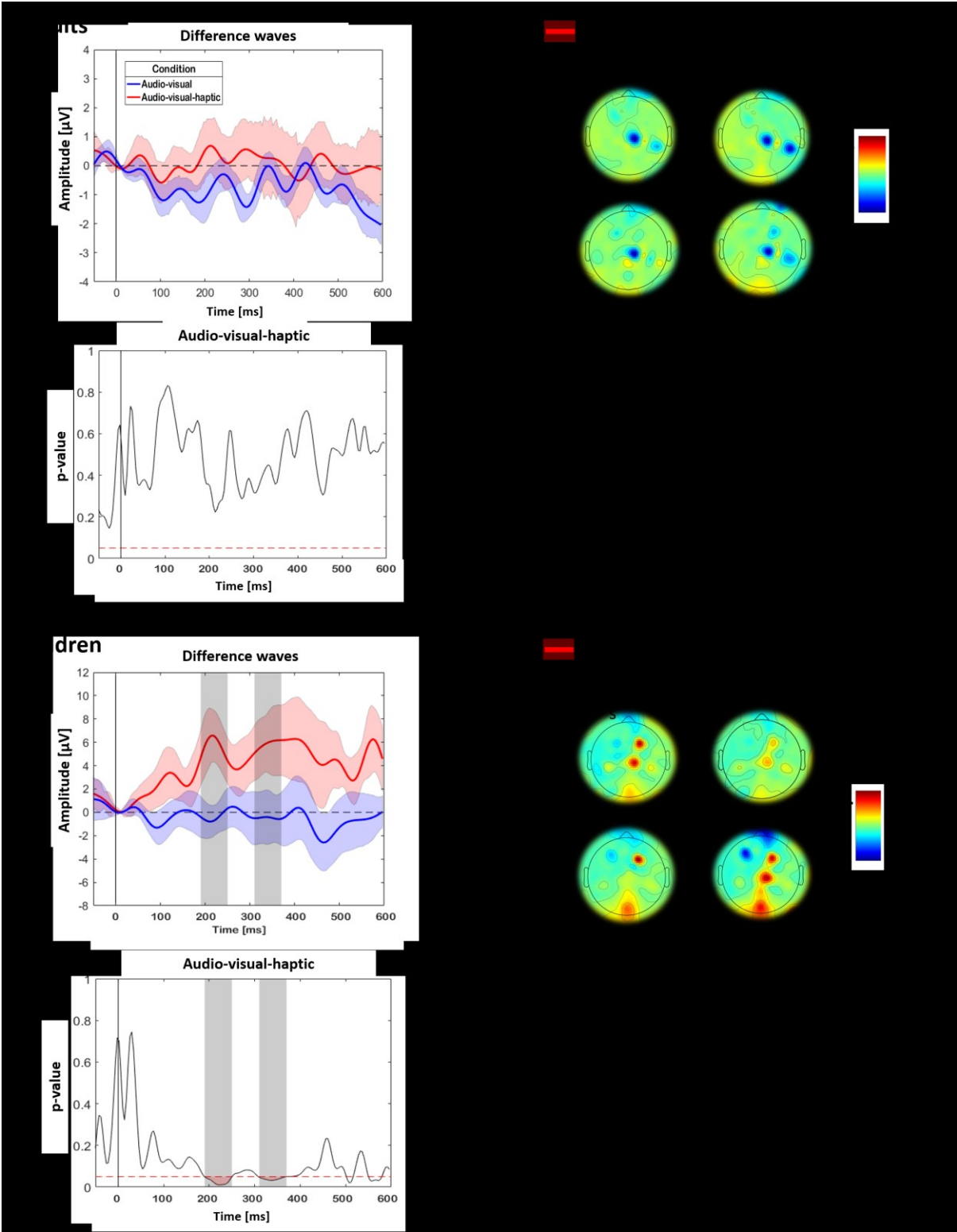


Figure 2 –

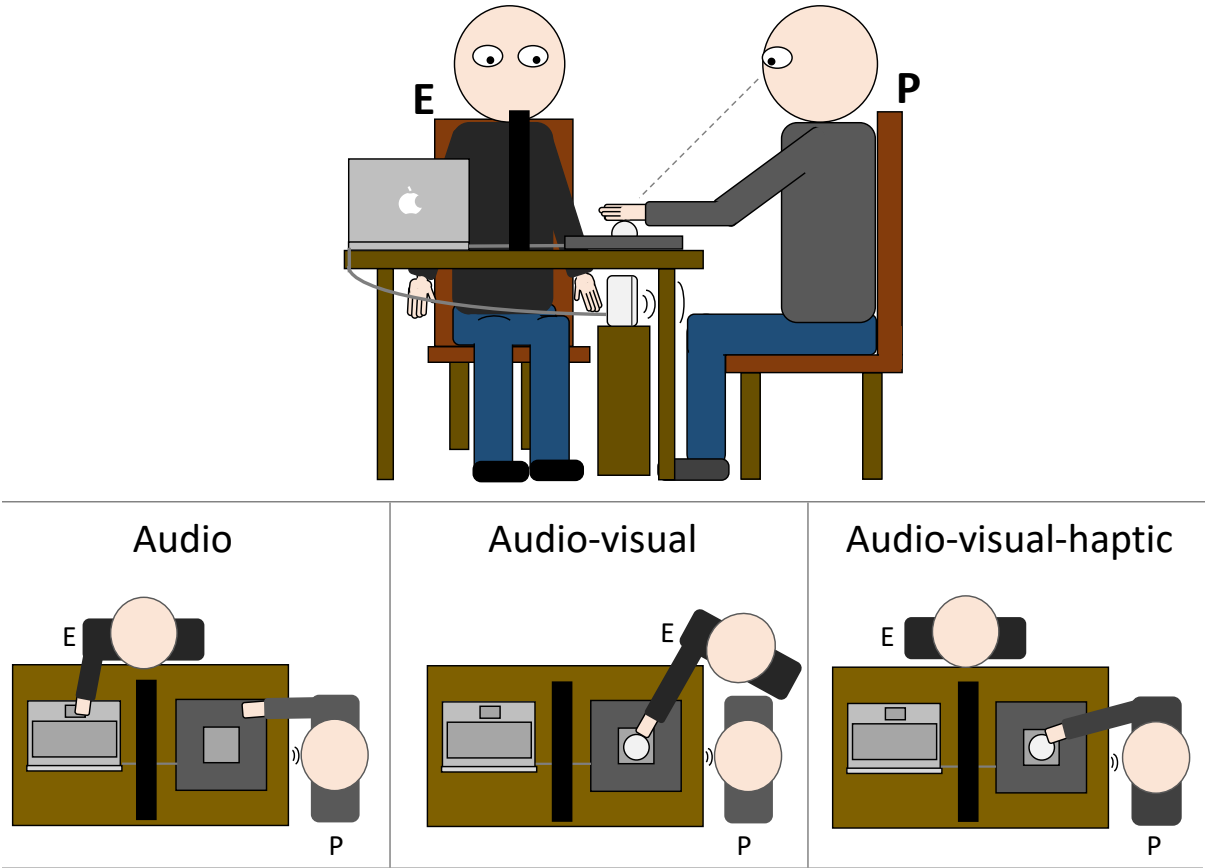


Figure 3 -

**Tables***Table 1.* Experimental design

<b>Block</b>	<b>Person tapping</b>	<b>Ball tapped</b>	<b>Auditory Stimuli</b>
Audio	None	None	small, big
Audio-visual	Experimenter	Small	small, big (congruent/incongruent)
Audio-visual-haptic	Participant	Small	small, big (congruent/incongruent)
Audio-visual	Experimenter	Big	small, big, (incongruent/congruent)
Audio-visual-haptic	Participant	Big	small, big (incongruent/congruent)